METHOD AND SYSTEM FOR REDUCING DEADLOCK IN FIBRE CHANNEL FABRICS USING VIRTUAL LANES

INVENTOR:

STEVEN M. BETKER

Cross Reference to related Applications

[0001] This application claims priority to U.S. provisional patent application serial number 60/542,241, filed on 02/05/2004, the disclosure of which is incorporated herein by reference in its entirety.

BACKGROUND

Field of the Invention

[0002] The present invention relates to Fibre Channel systems, and more particularly, to reducing deadlock problems in Fibre Channel Fabrics.

Background of the Invention

[0003] Fibre channel is a set of American National Standard Institute (ANSI) standards, which provide a serial transmission protocol for storage and network protocols such as HIPPI, SCSI, IP, ATM and others. Fibre channel provides an input/output interface to meet the requirements of both channel and network users.

[0004] Fibre channel supports three different topologies: point-to-point, arbitrated loop and fibre

channel fabric. The point-to-point topology attaches two devices directly. The arbitrated loop topology attaches devices in a loop. The fibre channel fabric topology attaches host systems directly to a fabric, which are then connected to multiple devices. The fibre channel fabric topology allows several media types to be interconnected.

[0005] Fibre channel is a closed system that relies on multiple ports to exchange information on attributes and characteristics to determine if the ports can operate together. If the ports can work together, they define the criteria under which they communicate.

[0006] In fibre channel, a path is established between two nodes where the path's primary task is to transport data from one point to another at high speed with low latency, performing only simple error detection in hardware.

[0007] Fibre channel fabric devices include a node port or "N_Port" that manages fabric connections. The N_port establishes a connection to a fabric element (e.g., a switch) having a fabric port or F_port. Fabric elements include the intelligence to handle routing, error detection, recovery, and similar management functions.

[0008] A fibre channel switch is a multi-port device where each port manages a simple point-to-point connection between itself and its attached system. Each port can be attached to a server, peripheral, I/O subsystem, bridge, hub, router, or even another switch. A switch receives messages from one port and automatically routes it to another port. Multiple calls or data transfers happen concurrently through the multi-port fibre channel switch.

[0009] Fibre channel switches use memory buffers to hold frames received and sent across a network. Associated with these buffers are credits, which are the number of frames that a buffer can hold per fabric port.

[0010] The following Fibre Channel standards are used for Fibre Channel systems and Fibre Channel Fabrics, and are incorporated herein by reference in their entirety:

[0011] ANSI INCITS xxx-200x Fibre Channel Framing and Signaling Interface (FC-FS) - T11/Project 1331D; and ANSI INCITS xxx-200x Fibre Channel Switch Fabric-3 (FC-SW-3), T11/Project 1508D

[0012] As discussed above, a Fibre Channel Fabric can consist of multiple switches connected in an arbitrary topology. The links between the switches use a

buffer-to-buffer credit scheme for flow control so that all frames transmitted have a receive buffer. Fabric deadlock may occur if a switch cannot forward frames because the recipient switch buffers (receive buffers) are full.

[0013] The following example, described with respect to Figure 6, shows how a deadlock situation can occur. Figure 6 shows five switches ("SW") 1, 2, 3, 4, and 5 that are linked together by ISLs (Inter Switch Links) in a ring topology. Host 11 and target 21 are linked to switch 1, host 12 and target 22 are linked to switch 2, and so forth.

[0014] In this example, hosts 11-15 can send data as fast as they can to a target that is two (2) hops away, for example:

Host 11 can send data to target 23;

Host 12 can send data to target 24;

Host 13 can send data to target 25;

Host 14 can send data to target 21; and

Host 15 can send data to target 22

[0015] For illustration purposes only, all traffic goes in the clockwise direction in Figure 6.

[0016] The receive buffers available for each ISL in the direction of traffic may get filled with frames addressed to the next switch.

[0017] For example:

For the ISL between switch 1 and switch 2, the receive buffers on switch 2 get filled with frames for switch 3;

For the ISL between switch 2 and switch 3, the receive buffers on switch 3 get filled with frames for switch 4;

For the ISL between switch 3 and switch 4, the receive buffers on switch 4 get filled with frames for switch 5;

For the ISL between switch 4 and 5,the receive buffers on 5 get filled with frames for switch 1; and

For the ISL between switch 5 and switch 1,the receive buffers on switch 1 get filled with frames for switch 2.

[0018] The transmit side of a switch waits for R_RDYs before it can transmit any frames. If frames cannot be transmitted from one ISL, then the receive buffers for the other ISL cannot be emptied. If the receive buffers cannot be emptied, no R_RDY flow control signals can be transmitted, which deadlocks the Fabric.

[0019] Many large Fabrics have paths that form rings within them, especially if they are designed to avoid single points of failure by using redundant switches. Such network traffic patterns may result in a deadlock situation disrupting networks using fibre channel switches and components.

[0020] Therefore, there is need for a system and method for minimizing deadlock problems in fibre channel switches.

SUMMARY OF THE PRESENT INVENTION

[0021] In one aspect of the present invention, a method for processing fibre channel frames is provided. The method includes, assigning a virtual lane for a frame based on a hop count for the frame; determining if the assigned virtual lane has available credit; and transmitting the frame if credit is available.

[0022] The method also includes, incrementing a counter value for counting available credit for the assigned virtual lane, if the frame is sent using the assigned virtual lane. If all credit for the assigned virtual lane has been used, then a next virtual lane is selected with non-zero credit.

[0023] In yet another aspect of the present invention, a method for processing fibre channel frames using a fabric switch element having a receive port and a transmit port is provided. The method includes assigning a virtual lane in the receive port based on a hop count for the frame; and sending a primitive to a transmit port with the assigned virtual lane.

[0024] The method further includes, assigning a virtual lane on the transmit port based on the hop count of the

frame; and determining if credit is available for the assigned virtual lane to send the frame. A credit count for the assigned virtual lane is maintained by a counter and the assigned virtual lane has a maximum credit count.

[0025] In another aspect of the present invention, a system for processing fibre channel frames is provided. The system includes, a fibre channel fabric switch element including a receive port for receiving fibre channel frames, which includes a look up table to assign a virtual lane based on a hop count of the frame; and a transmit port that receives a primitive with the assigned virtual lane by the receive port and the transmit port includes a credit control module that determines if an assigned virtual lane can transmit a frame based on available credit.

[0026] The credit control module increments a credit count for an assigned virtual lane if a frame has been transmitted from the assigned virtual lane. The credit control module also decrements a credit count for an assigned virtual lane if a VC_RDY is received. The credit control module also maintains a maximum count for every virtual lane used for transmitting frames. An increment selector is used to increment credit count and a decrement selector is used to decrease the credit

count. The credit control module also uses compare logic to compare available credit for an assigned virtual lane at any given time with a programmed maximum credit value for the assigned virtual lane.

[0027] In yet another aspect of the present invention, a system for processing fibre channel frames is provided. The system includes, the means for assigning dedicated virtual lanes for transmitting frames, where the virtual lanes are assigned based on a hop count of a frame; means for maintaining a credit count for each virtual lane used for transmitting frames; and means for determining if credit is available for a particular virtual lane that is assigned based on the hop count.

[0028] The system also includes the means for maintaining a maximum credit count for each virtual lane; and means for comparing the maximum credit count with the credit available for a virtual lane at any given time.

[0029] In yet another aspect of the present invention, a fibre channel fabric switch element for processing fibre channel frames, is provided. The switch element includes, means for assigning dedicated virtual lanes for transmitting frames, where the virtual lanes are assigned based on a hop count of a frame; means for

maintaining a credit count for each virtual lane used for transmitting frames; and means for determining if credit is available for a particular virtual lane that is assigned based on the hop count.

[0030] The switch element also includes, means for maintaining a maximum credit count for each virtual lane; and means for comparing the maximum credit count with the credit available for a virtual lane at any given time.

[0031] In yet another aspect, the present invention reduces/prevents the deadlock by separating frames queued for transmission into virtual lanes, each with its own transmit queue and flow control. Flow control uses the Fibre Channel VC_RDY primitive signal to give separate flow control signals to each virtual lane. Also, no frames are discarded to reduce/avoid deadlock.

[0032] This brief summary has been provided so that the nature of the invention may be understood quickly. A more complete understanding of the invention can be obtained by reference to the following detailed description of the preferred embodiments thereof in connection with the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0033] The foregoing features and other features of the present invention will now be described. In the

drawings, the same components have the same reference numerals. The illustrated embodiment is intended to illustrate, but not to limit the invention. The drawings include the following Figures:

[0034] Figure 1A is a block diagram of a fibre channel network;

[0035] Figures 1B -1D show block diagrams of fibre channel fabric switch elements, used according to one aspect of the present invention;

[0036] Figure 2 is a block diagram of a switch chassis used according to one aspect of the present invention;

[0037] Figure 3 shows a block diagram of a system in a receive port for assigning virtual lanes based on hop count, according to one aspect of the present invention;

[0038] Figure 4 is a block diagram showing a system in a transmit port for using virtual lanes based on hop count, according to one aspect of the present invention;

[0039] Figure 5 is a block diagram of a credit control module, used according to one aspect of the present invention;

[0040] Figure 6 shows a block diagram for illustrating deadlock in a fibre channel fabric;

[0041] Figure 7 is a flow diagram for using virtual lanes, according to one aspect of the present invention;

[0042] Figure 8 is a flow diagram for transmitting frames using virtual lane assignment, according to one aspect of the present invention;

[0043] Figure 9 is a flow diagram for VC_RDY processing, according to one aspect of the present invention;

[0044] Figure 10 is a flow diagram for processing frames at a receive port, according to one aspect of the present invention; and

[0045] Figure 11 is a flow diagram for handling frames in a transmit port, according to one aspect of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0046] Definitions:

[0047] The following definitions are provided as they are typically (but not exclusively) used in the fibre channel environment, implementing the various adaptive aspects of the present invention.

[0048] "D_ID": 24-bit fibre channel header field that contains destination address.

[0049] "EOF": End of Frame

[0050] "E-Port": A fabric expansion port that attaches to another Interconnect port to create an Inter-Switch Link.

[0051] "F-Port": A port to which non-loop N_Ports are attached to a fabric and does not include FL_ports.

[0052] "Fibre channel ANSI Standard": The standard (incorporated herein by reference in its entirety) describes the physical interface, transmission and signaling protocol of a high performance serial link for support of other high level protocols associated with IPI, SCSI, IP, ATM and others.

[0053] "FC-1": Fibre channel transmission protocol, which includes serial encoding, decoding and error control.

[0054] "FC-2": Fibre channel signaling protocol that includes frame structure and byte sequences.

[0055] "FC-3": Defines a set of fibre channel services that are common across plural ports of a node.

[0056] "FC-4": Provides mapping between lower levels of fibre channel, IPI and SCSI command sets, HIPPI data framing, IP and other upper level protocols.

[0057] "Fabric": The structure or organization of a group of switches, target and host devices (NL_Port, N_ports etc.).

[0058] "Fabric Topology": A topology where a device is directly attached to a fibre channel fabric that uses destination identifiers embedded in frame headers to route frames through a fibre channel fabric to a desired destination.

[0059] "FL_Port": A L_Port that is able to perform the function of a F_Port, attached via a link to one or more NL_Ports in an Arbitrated Loop topology.

[0060] "Inter-Switch Link" ("ISL"): A Link directly connecting the E_port of one switch to the E_port of another switch.

[0061] Port: A general reference to N. Sub.__--Port or F.Sub.__--Port.

[0062] "L_Port": A port that contains Arbitrated Loop functions associated with the Arbitrated Loop topology.

[0063] "N-Port": A direct fabric attached port.

[0064] "NL_Port": A L_Port that can perform the function of a N_Port.

[0065] "S_ID": 24 bit fibre channel header field that contains the source address of a frame.

[0066] "SOF": Start of Frame

[0067] "Switch": A fabric element conforming to the Fibre Channel Switch standards.

[0068] "VL" (Virtual Lane): A virtual portion of the data path between a source and destination port each having independent buffer to buffer flow control.

[0069] To facilitate an understanding of the preferred embodiment, the general architecture and operation of a fibre channel system will be described. The specific architecture and operation of the preferred embodiment will then be described with reference to the general architecture of the fibre channel system.

[0070] Figure 1A is a block diagram of a fibre channel system 100 implementing the methods and systems in accordance with the adaptive aspects of the present invention. System 100 includes plural devices that are interconnected. Each device includes one or more ports, classified as node ports (N_Ports), fabric ports (F_Ports), and expansion ports (E_Ports). Node ports may be located in a node device, e.g. server 103, disk array 105 and storage device 104. Fabric ports are located in fabric devices such as switch 101 and 102. Arbitrated loop 106 may be operationally coupled to switch 101 using arbitrated loop ports (FL_Ports).

[0071] The devices of Figure 1A are operationally coupled via "links" or "paths". A path may be established between two N_ports, e.g. between server

103 and storage 104. A packet-switched path may be established using multiple links, e.g. an N-Port in server 103 may establish a path with disk array 105 through switch 102.

[0072] Figure 1B is a block diagram of a 20-port ASIC fabric element according to one aspect of the present invention. Figure 1B provides the general architecture of a 20-channel switch chassis using the 20-port fabric element. Fabric element includes ASIC 20 with non-blocking fibre channel class 2 (connectionless, acknowledged) and class 3 (connectionless, unacknowledged) service between any ports. It is noteworthy that ASIC 20 may also be designed for other fibre channel classes of service, within the scope and operation of the present invention as described herein.

[0073] The fabric element of the present invention is presently implemented as a single CMOS ASIC, and for this reason the term "fabric element" and ASIC are used interchangeably to refer to the preferred embodiments in this specification. Although Figure 1B shows 20 ports, the present invention is not limited to any particular number of ports.

[0074] ASIC 20 has 20 ports numbered in Figure 1B as GL0 through GL19. These ports are generic to common

Fibre Channel port types, for example, F_Port, FL_Port and E-Port. In other words, depending upon what it is attached to, each GL_Port can function as any type of port. Also, the GL_Port may function as a special port useful in fabric element linking, as described below.

[0075] For illustration purposes only, all GL_Ports are drawn on the same side of ASIC 20 in Figure 1B. However, the ports may be located on both sides of ASIC 20 as shown in other figures. This does not imply any difference in port or ASIC design. Actual physical layout of the ports will depend on the physical layout of the ASIC.

[0076] Each port GL0-GL19 has transmit and receive connections to switch crossbar 50. One connection is through receive buffer 52, which functions to receive and temporarily hold a frame during a routing operation. The other connection is through a transmit buffer 54.

[0077] Switch crossbar 50 includes a number of switch crossbars for handling specific types of data and data flow control information. For illustration purposes only, switch crossbar 50 is shown as a single crossbar. Switch crossbar 50 is a connectionless crossbar (packet switch) of known conventional design, sized to connect 21 x 21 paths. This is to accommodate

20 GL_Ports plus a port for connection to a fabric controller, which may be external or internal to ASIC 20.

[0078] In the preferred embodiments of switch chasis described herein, the fabric controller is a firmware-programmed microprocessor, also referred to as the input/out processor ("IOP"). IOP 66 is shown in Figure 1C as a part of a switch chassis utilizing one or more of ASIC 20. As seen in Figure 1B, bi-directional connection to IOP 66 is routed through path 67, which connects internally to a control bus 60. Transmit buffer 56, receive buffer 58, control register 62 and Status register 64 connect to bus 60. Transmit buffer 56 and receive buffer 58 connect the internal connectionless switch crossbar 50 to IOP 66 so that it can source or sink frames.

[0079] Control register 62 receives and holds control information from IOP 66, so that IOP 66 can change characteristics or operating configuration of ASIC 20 by placing certain control words in register 62. IOP 66 can read status of ASIC 20 by monitoring various codes that are placed in status register 64 by monitoring circuits (not shown).

[0080] Figure 1C shows a 20-channel switch chassis S2 using ASIC 20 and IOP 66. S2 will also include

other elements, for example, a power supply (not shown). The 20 GL_Ports correspond to channel C0-C19. Each GL_Port has a serial/deserializer (SERDES) designated as S0-S19. Ideally, the SERDES functions are implemented on ASIC 20 for efficiency, but may alternatively be external to each GL_Port.

[0081] Each GL_Port may have an optical-electric converter, designated as OE0-OE19 connected with its SERDES through serial lines, for providing fibre optic input/output connections, as is well known in the high performance switch design. The converters connect to switch channels C0-C19. It is noteworthy that the ports can connect through copper paths or other means instead of optical-electric converters.

[0082] Figure 1D shows a block diagram of ASIC 20 with sixteen GL_Ports designated as GL0-GL15 and four 10G port control modules designated as XG0-XG3. ASIC 20 include a control port 62A that is coupled to IOP 66 through a PCI connection 66A.

[0083] IOP 66 is also shown in Figure 2 as a part of a switch chasis 201, containing switch ports 204, 207, 210 and 215. Each port as described above has a transmit port (segment), for example, 205, 208, 211 and 213, and receive port (segment), for example, 206, 209, 212 and 214.

[0084] Transmit ports and receive ports are connected by switch crossbar 50 so that they can transfer frames. IOP 66 controls and configures the switch ports.

[0085] In one aspect of the present invention, dividing frame traffic on ISLs into virtual lanes by assigning virtual lanes based on the number of hops to a destination switch, as described below reduces deadlock. Each virtual lane has its own buffer-to-buffer credit scheme. The term hop count means the number of ISLs a frame has to traverse before it reaches a destination switch.

[0086] To reduce and/or avoid deadlock in fibre channel switches, the following port requirements are used:

[0087] A receive port has receive buffers at least equal to the largest number of hops to a destination as seen by the transmit port of the switch that receives a frame. A receive buffer is also reserved for each hop count.

[0088] The hop count for frames can be derived from the data exchanged by switches using the standard FSPF protocol to set up routing within the Fabric, as described in the fibre channel standard, FC-SW-3, incorporated herein by reference in its entirety.

[0089] The transmit and receive ports assign a virtual lane to each hop count.

[0090] Each virtual lane is assigned some buffer-to-buffer credit. The total credit for all the virtual lanes is less than or equal to the total number of receive buffers available at a receive port.

[0091] Frames queued for transmissions are assigned a virtual lane and each hop count has a virtual lane.

[0092] Frames are transmitted if credit is available for its virtual lane.

[0093] Receive ports assign a virtual lane for the received frames. In one example, virtual lanes are assigned based on the hop count to a destination switch. If the destination is within the switch that received the frame, then the hop count is 0.

[0094] When a receive port empties a receive buffer, making it available for another frame, it sends a VC_RDY(n) primitive to the other end of the link. The VC_RDY contains the number of the virtual lane of the frame just processed. The fibre channel standard, FC-FS (incorporated herein by reference in its entirety) describes the VC_RDY primitive signal.

[0095] Using virtual lanes as discussed above, keeps the receive buffers of an ISL from filling up with

frames for the same destination and provides receive buffers space for frames to other destination.

[0096] The following sequence shows how frames are delivered, using one aspect of the present invention:

[0097] Frames arriving on an ISL addressed to a particular switch can empty its receive buffers for virtual lane 0 and return VC_RDY(0) to the sender.

[0098] A switch that is 1 hop away from a destination switch (for example, switch 2 if the destination switch is switch 3 for a frame sent from switch 1) gets VC_RDY(0). In this example, this will allow switch 2 to empty the receive buffer for frames received from other switches.

[0099] During the exchange of ELP messages (fibre channel standard log-in messages) to log in the ISL ports (per FC-SW-3 standard, incorporated herein by reference in its entirety) virtual lanes and hop counts in the ELP messages can be used for flow control. If both ends of the ISL agree, the virtual lane flow control option is used. Since the virtual lane assignment is derived from the hop count, there is no need to negotiate virtual lane assignments.

[0100] If the receive port on an ISL receives a frame that it cannot route, it returns a VC_RDY(255) to the sender after the receive buffer is emptied. If a

VC_RDY(255) is received, the credit is allocated to the lowest numbered virtual lane that does not have maximum transmit credit available. If any VC_RDY(n) is received where virtual lane n does not exist or is already at maximum transmit credit, the credit is allocated the same way.

[0101] Figure 3 shows how received frames are processed using Virtual Lanes, according to one aspect of the present invention. The Fibre Channel header D_ID field 301A and a look-up table in logic 302 is used to look up the hop count for a frame based on the destination. The hop count number is used to assign a Virtual Lane (302A). The look-up table in logic 302 is loaded with data derived from the standard "FSPF" routing algorithm to determine the hop count. The frame and the assigned Virtual Lane 302A are stored in receive buffers 301.

[0102] When a frame (301A) is moved out of a receive buffer to a transmit port, receive buffer 301 sends a signal 303 to the transmit port. If the port is an ISL using the deadlock prevention process of the present invention, the flow control signal 303 is a VC_RDY primitive containing the assigned virtual lane (302A)when the frame was received.

[0103] Figure 4 shows a block diagram of logic 400 used in flow control for a transmit port using virtual Lanes (in this example, 4 virtual lanes) for reducing and/or avoiding deadlocks, according to one aspect of the present invention. Logic 400 can be used for ISLs (E-ports) that connect switches. Frames are queued in queue 401 and are assigned a Virtual Lane (401A).

[0104] In one aspect of the present invention, the assigned Virtual Lane 401A is one less than the Virtual Lane assigned by the receive port (303, Figure 3). The assigned Virtual Lane 401A is used by multiplexer 403 to determine which credit control output from credit control module 402 is used to determine if a frame has available credit for transmission. Transmit control module 404 determines if a frame is available for transmission and credit is available. If a frame and credit is available, then frame 405 is transmitted to its destination.

[0105] Figure 5 shows a block diagram of credit control module 402, according to one aspect of the present invention. Credit control module 402 has an increment selector 501 and decrement selector 502 for increasing and decreasing counters 503. Counter 503 maintains count for each virtual lane, for example, VL0, VL1, VL2 and VL3. Selector 501 increments a

virtual lane credit count based on 501B, i.e., the VL of a frame that is transmitted. Selector 501 also receives input 501A from transmit control module 404 which indicates when a frame has been sent.

[0106] If a VC_RDY 502A is received from a receive buffer, then selector 502 decrements the value of the appropriate counter 503. Selector 502 also receives the Virtual Lane associated with the VC_RDY (502) from the receive buffer (similar to 303, Figure 3).

[0107] Counter(s) 503 maintain(s) count for the virtual lanes VL0, VL1, VL2 and VL3. In one aspect of the present invention, each virtual lane may have a pre-programmed maximum count value that is stored in counters 504.

[0108] Logic (also referred to as "compare module") 505 compares the maximum count value for virtual lane 0 to determine if credit is available on virtual lane 0. Logic 506-508 performs the same function for virtual lanes 1, 2 and 3, respectively. Compare modules 505-508 generate signals 505A-505D indicating if credit is available for a particular Virtual Lane to transmit a frame.

[0109] To illustrate the adaptive aspects of the present invention separate counters have been shown, however, the present invention is not limited to any

particular number of counters. For example, logic with a single counter may be used to compare the maximum count (504) and the count (503) for each lane.

[0110] In another aspect of this invention, other queuing methods could be used instead of the one described for this embodiment. For instance, a transmit port may have a transmit queue for each Virtual Lane, and/or for each receive port.

[0111] Figure 7 shows a process flow diagram, according to one aspect of the present invention using virtual lanes for frame transmission.

[0112] In step S700, the process selects a particular virtual lane, for example, virtual lane 0. In step S701, the process determines if the credit used for a particular virtual lane (for example, virtual lane 0) is less than a maximum programmed amount or a particular value (Figure 5, 504). This is performed by logic 505-508, depending upon which lane is selected. In the foregoing example, for VL0, it is logic 505. If virtual lane credit exceeds the maximum count, then in step S705, the process selects the next available virtual lane (for example, virtual lane 1). In step S706, the process determines if the selected virtual lane has used it's maximum credit. If yes, the process

reverts back to step S700, otherwise the process moves to step S701.

[0113] In step S702, the process determines if a frame is available for the selected virtual lane (i.e. the lane that is selected in step S701 or S706). If a frame is not available, the process moves to step S705.

[0114] If a frame is available in step S702, the frame is sent in step S703 (for example, 505A), and then in step S704, the credit counter for the virtual lane is incremented by selector 501 using one of the counters 503.

[0115] Figure 8 provides a flow diagram for selecting a virtual lane, according to one aspect of the present invention. In step S800, the process determines the hop count for a frame's D_ID. In step S801, the process sets the virtual lane of a frame (401A). In one example, the virtual lane is one less than the hop count.

[0116] Figure 9 shows a flow diagram for processing VC_RDY primitives, according to one aspect of the present invention. In step S900, the process uses a VC_RDY virtual lane number to index virtual lane credit that has been used for a particular frame. This information comes as 502B to selector 502 from the receive port.

[0117] In step S901, the process determines if all the credit for a particular virtual lane has been used. If yes, then in step S903, the process finds the next virtual lane with non-zero credit.

[0118] If in step S901, virtual lane credit is not equal to zero, then in step S902, selector 502 decrements the virtual lane credit value that has been used.

[0119] Figure 10 shows a flow diagram for processing frame in the receive port of a switch, according to one aspect of the present invention. In step S1000, the process receives frame D-ID 301A. In step S1001, the process determines the hop count for the frame. This can be obtained by using the standard FSPF algorithm.

[0120] In step S1002, the process assigns a virtual lane based on the hop count. If the frame is destined for the same switch, the virtual lane is zero.

[0121] In step S1003, the receive port sends the frame to the transmit segment and in step S1004, a VC_RDY primitive is sent to the transmit port with the VL assignment value (502B).

[0122] Figure 11 is an overall process flow diagram for sending frames from the transmit port of a switch. In step 1100, the process assigns a virtual lane for a frame that is queued (401A).

[0123] In step S1101, the process determines if credit is available for a particular virtual lane. This is performed by credit control module 402, as described above.

[0124] In step S1102, a frame is sent if credit is available.

[0125] In one aspect, the present invention reduces/prevents the deadlock by separating frames queued for transmission into virtual lanes, each with its own transmit queue and flow control. Flow control uses the Fibre Channel VC_RDY primitive signal to give separate flow control signals to each virtual lane. Also, no frames are discarded to reduce/avoid deadlock.

[0126] Although the present invention has been described with reference to specific embodiments, these embodiments are illustrative only and not limiting. Many other applications and embodiments of the present invention will be apparent in light of this disclosure and the following claims.